

0590
0731

10



ENTERED

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/855,340A

DATE: 08/07/2002

TIME: 10:53:29

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\08072002\I855340A.raw

3 <110> APPLICANT: Hosted, Jr., Thomas J.
 4 Horan, Ann C.
 6 <120> TITLE OF INVENTION: Isolation of Micromonospora carbonacea var africana
 7 pMLP1 integrase and use of integrating function for
 8 site-specific integration into Micromonospora
 9 halophilica and Micromonospora carbonacea chromosome
 11 <130> FILE REFERENCE: IN01164K
 13 <140> CURRENT APPLICATION NUMBER: 09/855,340A
 C--> 14 <141> CURRENT FILING DATE: 2001-05-15
 16 <150> PRIOR APPLICATION NUMBER: 60/204,670
 17 <151> PRIOR FILING DATE: 2000-05-17
 19 <160> NUMBER OF SEQ ID NOS: 16
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 1179
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Micromonospora carbonacea
 28 <400> SEQUENCE: 1
 29 gtgtggatcg agaagaacgg gcccgtctac cgcatcgaaa acctcggttcg cggtaaaaag 60
 30 gtcaccatc agaccgtta tccgacgaag accagcgcca agaatgcgt ggtcagttc 120
 31 cgtgcggagc agttgcaggg caacgcgtc atgcgcgcg gcggtcagat taccctcgcc 180
 32 gatttcgtgg gggagtggg gccgagctac gaaaagacgc taaaaaccgac cgccgtgaac 240
 33 tcggagggca accggatccg caaccacctc ctgcccatac tcggccatct cacccttgac 300
 34 gagctggacg ggcagggtcac ccacgcgtgg gtcaacgacc tggaggccgg cgtcggcccg 360
 35 tggccggagt ccacgcgggg tgcgcggaa acgcgttgcg cgaagacgt cagcaactgc 420
 36 cacggcctgc tgcacacgt ctgcgcgcg gcatgcgcg cggaaacggat caggctcaac 480
 37 ccgtgcttt cgacgtatgtc gccccggcgc gagccgaaag agatgaagtt cctgagcgc 540
 38 ccggagatcg gtccggcttat cacggcgctt ccgcgcact ggccggcgct cgtcatgt 600
 39 ctgtggcga ccggcttgag gtgggggtgag gcatgcgcg tgcgcgcgg ccgggtcgac 660
 40 ctgctcgccg cgccggcccg gctgaccgtc gtcgagcgc tccaggagct ggccacgc 720
 41 ggagagatcg tcttccagtc gccaagagacc gcaaggggcc ggccgcacggt cagtttacc 780
 42 acgaaagtgc ctctactgtc tacgcactc atgcgcggaa agaaaaagtga cgaggctgt 840
 43 ttcacccgcg cggaaaggcg gatggtaagg acgcgcattt tccggcgat ctgggtcaag 900
 44 gcgtgcgagg aagccgggtc tccggctta cgcattcactc atctgcggca cactcacgc 960
 45 gcatgcgttca tttctgcgg ggtccgcgtc tcggcgatct cccggccgcg cggtcactcg 1020
 46 tcgatgcgcg tcacggatct gctgtacggg cacctgcgtc aggagggtcga cgaggggatc 1080
 47 ctgcggcga tcgaggaggc gatggccggc gtccgggtcg aggacctgg a ggcggactc 1140
 48 gacgaggagc tgacggacgt gttggccgac gcagcatga 1179
 51 <210> SEQ ID NO: 2
 52 <211> LENGTH: 426
 53 <212> TYPE: DNA
 54 <213> ORGANISM: Micromonospora carbonacea
 56 <400> SEQUENCE: 2

RAW SEQUENCE LISTING DATE: 08/07/2002
 PATENT APPLICATION: US/09/855,340A TIME: 10:53:29

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\08072002\I855340A.raw

57 atgcgcaaca caccggggct ggggcgcggc acatggccg catacgtcct caccgcccgc 60
 58 gagcgcgcgg gactgaccaa gagcggatgg gccagcgcga tccagaaggaa ccggccacc 120
 59 gtccggccgg gggaggacgg caagaaccgg cccgacgacg cggacctcggt tgcccgccgtc 180
 60 gcccagggtgc tcggcctcga cctcgacgaa gcccctcgccg ccgcagggtct gcgcggccgg 240
 61 gtcaccccgcc cagcgacccc aaccatggac ctggacgagg aaatcgagct ggtccgcacc 300
 62 gaccccaagc tggacgagga catgaagcggc cgcatcatcg ccctaattcct ggagcgcgt 360
 63 gagcgcgaca aggccggccg gatcgaggaa accaagcggc tcatcgaccc gttccggccgg 420
 64 agctga 426
 67 <210> SEQ ID NO: 3
 68 <211> LENGTH: 34
 69 <212> TYPE: DNA
 70 <213> ORGANISM: Micromonospora carbonacea
 72 <400> SEQUENCE: 3
 73 ccccggtacg ggttcaattc ccatcagtca ccccg 34
 76 <210> SEQ ID NO: 4
 77 <211> LENGTH: 241
 78 <212> TYPE: DNA
 79 <213> ORGANISM: Micromonospora carbonacea
 81 <400> SEQUENCE: 4
 82 tattagtccg cacgcccggcc ggcggccgg gaggcggagcg catggtggt gtagctcagt 60
 83 tggcagagca cccgttgggtg gtcccggttg tcgtgggttc aattcccatc agtcacccgt 120
 84 acacgaaggc cccctccact cggaggggggc ctgcggcggtt cctgagggtt cgcggtcagg 180
 85 cggcggctc ggcgttgggg gactcggccc cgtcggcggtt agtggcctcg gcgtccgggg 240
 86 a 241
 89 <210> SEQ ID NO: 5
 90 <211> LENGTH: 243
 91 <212> TYPE: DNA
 92 <213> ORGANISM: Micromonospora carbonacea
 94 <400> SEQUENCE: 5
 95 tggcggggggt gtggcttata ttatcgca cggccggccgg ccccgccggc gcggagcgcga 60
 96 tggcggtgt agctcgttg cagagcacc gggttgtgtt cccgttgcgtt gttttcaaa 120
 97 ttcccatcag tcacccggca agtggatcta ctccacagca gatcaggccc cctccgaaga 180
 98 gggggcctga tgcgtcatag gggacaggtt gggaaactca accccggctt cttgcgtcgc 240
 99 gtc 243
 102 <210> SEQ ID NO: 6
 103 <211> LENGTH: 247
 104 <212> TYPE: DNA
 105 <213> ORGANISM: Micromonospora carbonacea
 107 <400> SEQUENCE: 6
 108 taggggaatc cactccggag acgcccggag caatccggag catgacggag caaccagcag 60
 109 gtcagggtggc ctgttaccc cctgaccagg gcccgggtac gggttcaatt cccatcgtc 120
 110 acccgtacac gaaggcccccc tccactcgga gggggccttc ggcgttccgtt agggttcg 180
 111 gtcaggcggtt cggctcgccg ctggggactt cggccgggtc ggcggagtg gcctcgccgt 240
 112 cccggga 247
 115 <210> SEQ ID NO: 7
 116 <211> LENGTH: 255
 117 <212> TYPE: DNA
 118 <213> ORGANISM: Micromonospora halophytica
 120 <400> SEQUENCE: 7

RAW SEQUENCE LISTING DATE: 08/07/2002
 PATENT APPLICATION: US/09/855,340A TIME: 10:53:29

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\08072002\I855340A.raw

121 tttctccgca cccgccccggg gcgttcgacc ggggtcgccg gcatgggtgc ttagctcg 60
 122 ttggcagagc accgggttgt gtcgggtt caattccat cagtcacccc 120
 123 agtaagacc caggtcaggg ccgttctca ccggccctga cgcatttca gggcatgg 180
 124 gggggcgcta ccgggggtgg ggtgtctcac cgcgagccag catctcgatc aggcgatcg 240
 125 gccggcgctg ccggg 255
 128 <210> SEQ ID NO: 8
 129 <211> LENGTH: 315
 130 <212> TYPE: DNA
 131 <213> ORGANISM: Micromonospora halophytica
 133 <400> SEQUENCE: 8
 134 tttctccgca cccgccccggg gcgttcgacc ggggtcgccg gcatgggtgc ttagctcg 60
 135 ttggcagagc accgggttgt gtcgggtt caattccat cagtcacccc 120
 136 gcaagtggat ctactccaca gcagatcagg cccctccga agagggggcc ttagatcg 180
 137 taggggacac gtaggggaaac tcaacccccc gtccttgct cgcgtcggtt catgccgtcc 240
 138 gcttacccct ccgcgtaccc ggcctctcc cgttccctga tctcggcg 300
 139 cgcagggtcg cctcc 315
 142 <210> SEQ ID NO: 9
 143 <211> LENGTH: 260
 144 <212> TYPE: DNA
 145 <213> ORGANISM: Micromonospora halophytica
 147 <400> SEQUENCE: 9
 148 taggggaatc cactccggag acgccccggag caatccggag catgacggag caaccagg 60
 149 gtcaggtggc ctgttaccc cctgaccagg gccccggtaat gggttcaatt cccatcagtc 120
 150 accccaggtt agaccaggat cagggccggc tctcaccggc cctgacgcat tttcaggggc 180
 151 atgggtgggg cgctaccggg ggtgggtgtt ctcaccggc ggcagcatct cgatcaggcg 240
 152 atcgagccgg cgctccggg 260
 154 <210> SEQ ID NO: 10
 155 <211> LENGTH: 209
 156 <212> TYPE: DNA
 157 <213> ORGANISM: artificial sequence
 159 <220> FEATURE:
 160 <223> OTHER INFORMATION: pMLP1 attP region
 162 <400> SEQUENCE: 10
 163 taggggaatc cactccggag acgccccggag caatccggag catgacggag caaccagg 60
 165 gtcaggtggc ctgttaccc cctgaccagg gccccggtaat gggttcaatt cccatcagtc 120
 167 accccggcaag tggatctact ccacaggcaga tcaggcccc tccgaagagg gggcctgatg 180
 169 cgtcataggg gacaggtagg ggaactcaa 209
 172 <210> SEQ ID NO: 11
 173 <211> LENGTH: 19
 174 <212> TYPE: DNA
 175 <213> ORGANISM: artificial sequence
 177 <220> FEATURE:
 178 <223> OTHER INFORMATION: primer PR144
 180 <400> SEQUENCE: 11
 181 tgcttcgacg ccatcargg 19
 184 <210> SEQ ID NO: 12
 185 <211> LENGTH: 20
 186 <212> TYPE: DNA
 187 <213> ORGANISM: artificial sequence

RAW SEQUENCE LISTING DATE: 08/07/2002
 PATENT APPLICATION: US/09/855,340A TIME: 10:53:29

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\08072002\I855340A.raw

189 <220> FEATURE:
 190 <223> OTHER INFORMATION: primer PR145
 192 <220> FEATURE:
 193 <221> NAME/KEY: misc_feature
 194 <222> LOCATION: (7)..(7)
 195 <223> OTHER INFORMATION: n is inosine (I)
 198 <400> SEQUENCE: 12
 W--> 199 **gttggaaacctt ccggaaatccgc** 20
 201 <210> SEQ ID NO: 13
 202 <211> LENGTH: 20
 203 <212> TYPE: DNA
 204 <213> ORGANISM: artificial sequence
 206 <220> FEATURE:
 207 <223> OTHER INFORMATION: primer PDH504
 209 <400> SEQUENCE: 13
 210 **aggggcaacaa gggaaacgtc** 20
 213 <210> SEQ ID NO: 14
 214 <211> LENGTH: 21
 215 <212> TYPE: DNA
 216 <213> ORGANISM: artificial sequence
 218 <220> FEATURE:
 219 <223> OTHER INFORMATION: primer PDH505
 221 <400> SEQUENCE: 14
 222 **ggcgccgggtt tggctattat t** 21
 225 <210> SEQ ID NO: 15
 226 <211> LENGTH: 21
 227 <212> TYPE: PRT
 228 <213> ORGANISM: artificial sequence
 230 <220> FEATURE:
 231 <223> OTHER INFORMATION: amino acid sequence of open reading frame indicated in
 figures 4b
 232 and 4d
 234 <400> SEQUENCE: 15
 236 Ser Pro Asp Ala Glu Ala Thr Pro Ala Asp Gly Ala Glu Ser Pro Ser
 237 1 5 10 15
 240 Ala Glu Pro Thr Ala
 241 20
 244 <210> SEQ ID NO: 16
 245 <211> LENGTH: 21
 246 <212> TYPE: PRT
 247 <213> ORGANISM: artificial sequence
 249 <220> FEATURE:
 250 <223> OTHER INFORMATION: amino acid sequence of open reading frame indicated in
 figures 5b
 251 and 5d
 253 <400> SEQUENCE: 16
 255 Arg Gln Arg Arg Leu Asp Arg Leu Ile Glu Met Leu Ala Arg Gly Glu
 256 1 5 10 15
 259 Thr Pro His Pro Arg
 260 20

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 08/07/2002
PATENT APPLICATION: US/09/855,340A TIME: 10:53:30

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08072002\I855340A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 7

VERIFICATION SUMMARY DATE: 08/07/2002
PATENT APPLICATION: US/09/855,340A TIME: 10:53:30

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08072002\I855340A.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:199 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0